The coevolution of cooperation and cognition in humans

Miguel dos Santos[1,*] and Stuart A. West[1,†]

[1]Department of Zoology, University of Oxford, UK

May 2, 2018

Abstract

Cooperative behaviours in archaic hunter-gatherers could have been maintained partly due to the gains from cooperation being shared with kin. However, the question arises as to how cooperation was maintained after early humans transitioned to larger groups of unrelated individuals. We hypothesize that after cooperation had evolved via benefits to kin, the consecutive evolution of cognition increased the returns from cooperating, to the point where benefits to self were sufficient for cooperation to remain stable when group size increased and relatedness decreased. We investigate the theoretical plausibility of this hypothesis, with both analytical modeling and simulations. We examine situations where cognition either (i) increases the benefits of cooperation, (ii) leads to synergistic benefits between cognitively enhanced cooperators, (iii) allows the exploitation of less intelligent partners, and (iv) the combination of these effects. We find that cooperation and cognition can coevolve - cooperation initially evolves, favouring enhanced cognition, which favours enhanced cooperation, and stabilizes cooperation against a drop in relatedness. These results suggest that enhanced cognition could have transformed the nature of cooperative dilemmas faced by early humans, thereby explaining the maintenance of cooperation between unrelated partners.

*miguel.dossantos@zoo.ox.ac.uk

†stuart.west@zoo.ox.ac.uk

1 Introduction

Hunting and gathering was the main subsistence strategy of archaic humans over the last 2 million years, until the advent of agriculture 10 thousand years ago (Barker, 2009; Bocquet-Appel, 2011). Modern hunter-gatherer societies live in small bands mainly composed of unrelated individuals (Hill et al., 2011; Marlowe, 2005). However, the first hunter-gatherers of the genus *Homo*, more than 1.5 million years ago, lived in smaller groups where partners were probably more related than in modern hunter-gatherers (Aiello and Dunbar, 1993; Bittles and Black, 2010; Dunbar, 2009; Hatala et al., 2016; Lalueza-Fox et al., 2011). The higher relatedness in archaic hunter-gatherers could explain, at least partially, a range of cooperative behaviours, such as group hunting and meat sharing, because the benefits of cooperation were likely to be shared with relatives (kin selection; Hamilton (1964)). The problem is to explain how a similarly high level of cooperation can be preserved during the transition to larger groups of unrelated individuals.

Our hypothesis is that after cooperation had evolved in small groups, via kin selection, coevolution with another trait increased the returns from cooperation, allowing it to be maintained when group size increased and relatedness decreased. In particular, we suggest that once cooperation had evolved, natural selection favoured additional traits, such as enhanced cognition, which increased the efficiency of cooperative behaviours, or even enabled synergistic effects between individuals involved in cooperative actions, such that cognitively enhanced cooperators were able to produce larger benefits than without cognition. Examples of such potential traits range from increased

relative brain size, changes in brain connectivity and functionality, to genes involved in language (Andres et al., 2004; Ferland et al., 2004; Krause et al., 2007; Rightmire, 2004; Schoenemann, 2006). These innovations are believed to have augmented the social and technical intelligence of early humans, and might have been the basis for stone tool technologies, hunting, as well as the ability to imagine future outcomes, take the mental perspective of others (Byrne and Whiten, 1989; Dunbar, 1998, 2009; Buss, 2015; Somel et al., 2013; Stout et al., 2008, 2015; Tomasello et al., 2005; Whiten and Erdal, 2012; Vallender et al., 2008; Yeshurun et al., 2007; Byrne and Whiten, 1989; Whiten and Byrne, 1997). These enhanced cognitive capacities could have transformed the nature of some cooperative dilemmas faced by early humans, such that cooperation could still be stable, with low or negligible relatedness.

However, it is not clear whether cognition can coevolve with cooperation in the way that would be required by our hypothesis. While sociality is at the center of most explanations for the evolution of cognition (the 'social brain hypothesis'), the role of relatedness between social partners has often been argued to be of minor importance, at least in the later stages of human evolution (Alexander, 1990; Dunbar, 2009; Tomasello et al., 2012; Whiten and Erdal, 2012). Theoretical studies have therefore focused on situations where conflict, either within- or between-groups was the underlying factor promoting higher cognition (Byrne and Whiten, 1989; McNally et al., 2012; McNally and Jackson, 2013; Gavrilets, 2015). Furthermore, higher cognitive abilities could enable individuals to generate larger benefits for their group and, at the same time, take advantage of their social partners through de-

ception Byrne and Whiten (1989); Whiten and Byrne (1997). Hence, it is still unknown what the relationships are between relatedness and the benefits from both cooperation and cognition, and under what conditions both traits can coevolve and remain stable.

We test the theoretical plausibility of our hypothesis, by examining when cognition can coevolve with cooperation in groups of related individuals, and whether this coevolution can stabilise cooperation, even if subsequently relatedness between social partners decreases. In an analytical model, we consider different scenarios where cognition allows individuals to (i) gain greater benefits from cooperation by enabling them to either generate larger gains for the same cost (e.g. by predicting a prey's reactions), (ii) generate synergistic benefits as the number of cognitively enhanced individuals increases (e.g., through better coordination and/or communication with each other Boesch (2002); Tomasello et al. (2005); Call (2009); Brosnan et al. (2010)), and (iii) exploit less intelligent partners through manipulation or deception (Byrne and Whiten, 1989; Whiten and Byrne, 1997). Second, we ask whether some level of cooperation can be maintained once a cognitive trait has evolved, even though group size increases and relatedness decreases. We confirmed the robustness of our analytical results with individual-based simulations. Although we focus on human cognition, we stress that our theory applies more generally to other intelligent species, such as primates and cetaceans, where cognition is likely to have evolved in groups of related individuals (Boesch, 2002; Smith et al., 1981; Gazda et al., 2005; Pruetz and Bertolani, 2007).

2 Model

2.1 Model description

We consider an infinitely large population, which is subdivided into an infinite number of patches of size n (Wright, 1931). Individuals are haploid, and interact socially within patches. We assume that social interactions affect an individual's fecundity. After the social interactions, adults on patches produce a very large number of juveniles, and die. Juveniles all disperse to some new patch. Competition between juveniles reduces patch size to n individuals. We assume that individuals on each patch can be related, although we do not yet specify how relatedness comes about (e.g., juveniles might not disperse independently). Generations are non-overlapping and competition is global. We later expand our model to include limited dispersal and overlapping generations, and let relatedness vary in terms of the parameters of the life cycle.

Individuals carry two social traits: (i) a cooperative trait x, determining the probability of contributing, at personal fecundity cost $c > 0$, a baseline public good $b > 0$ that is shared equally among all group members, and (ii) a cognitive trait y, determining the probability of investing into better cognition capabilities early in development, at personal cost $d > 0$. We assume that both traits can be expressed independently, e.g., y can be expressed even though x is not, and we also assume no genetic correlation between the two traits, such that a change in the value of one trait does not influence that of the other trait.

We assume that cognition can enhance the contribution of a focal individual in two different ways. First, the contribution of individuals expressing y might merely generate better returns compared to the baseline contribution b. In this case, the benefit generated is increased by an amount $b_C \geq 0$. Second, individuals expressing the y trait might interact together in order to produce increasingly large public goods, i.e., synergy could occur between cognitively enhanced individuals, with a parameter $\alpha \geq 0$ controlling synergistic effects. Specifically, synergy occurs when $\alpha > 0$. Importantly, all individuals in the group enjoy an equal share of the total contributions of both normal and cognitively enhanced individuals.

Following from our assumptions, the amount of public good received by a focal individual takes the form $B(x_g, y_g) = x_g(1 - y_g)b + x_g y_g(b + b_C)\mathrm{e}^{\alpha y_g}$, where x_g and y_g are the group average cooperation and cognition traits, respectively, in the focal individual's group (including itself). Depending on the parameters b_C and α, this production function covers scenarios in which cognition may allow for (i) larger contributions, (ii) synergism between cognitively enhanced contributors, or (iii) both (figure 1).

These assumptions reflect a type of public good that is rivalrous because the per capita share depends on the number of consumers in the group (i.e. patch size n). However, our model can also reflect a non-rivalrous public good (e.g. cultural knowledge, such as the ability to create a tool) by multiplying the per capita benefit $B(x_g, y_g)$ by n, so that individuals receive a public good which only depends on the number of contributors in the group. As a consequence, in the conditions for cooperation and cognition to be stable

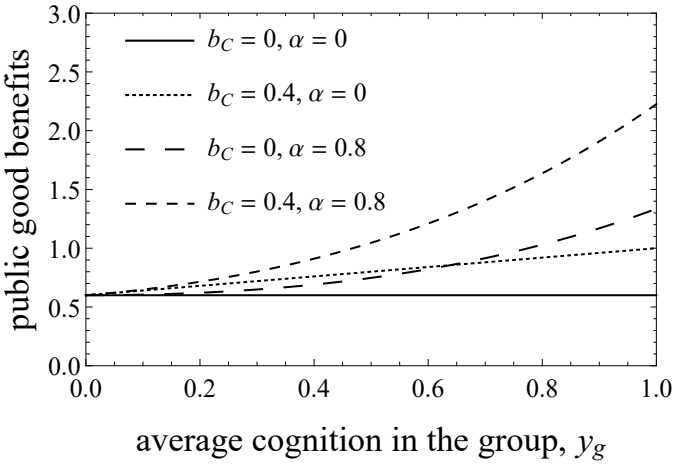

Figure 1: Public good production functions. Varying both b_C and α allows us to capture different potential effects that cognition might have on the per capita benefit from a public good that is rivalrous $B(x_g, y_g)$. In case of a non-rivalrous public good, the per capita benefit received by a focal individual is $nB(x_g, y_g)$. Parameters: $x_g = 1$, $b = 0.6$.

can be recovered from Tables 1 and S1-2, except that the benefit from the public good need to be multiplied by n.

With our assumptions, the fitness w of a focal individual is given by $w = F/\bar{F}$. Here, $F = 1 - x_0 c - y_0 d + B(x_g, y_g)$ is the fecundity of the focal individual, where x_0 and y_0 are the focal's cooperation and cognition traits, respectively, and $\bar{F} = 1 - \bar{x}c - \bar{y}d + B(\bar{x}, \bar{y})$ is the average fecundity in the population.

We follow the approach on the joint evolution of multiple social traits of Brown and Taylor (2010). Specifically, we consider the successive invasion of mutants in one trait in a resident population that is monomorphic for both traits, and where a mutant for one of the two traits (holding the other constant) differs only slightly from that of the resident population. We assume that a mutant will be rare globally, but potentially common in the local

group, due to relatedness.

We determine Hamilton's selection gradient for both traits to investigate when an increase in either trait is favoured by selection (Taylor and Frank, 1996; Frank, 1998). We denote R the relatedness of the focal individual to a random group member, including itself (i.e., 'whole-group' relatedness). For simplicity, we assume that relatedness is the same at both loci. Hence, the marginal inclusive fitness effects for cooperation and cognition are given by $H(x,y) = \partial w/\partial x_0 + R\partial w/\partial x_g$ and $K(x,y) = \partial w/\partial y_0 + R\partial w/\partial y_g$ respectively, where all derivatives are evaluated at $x_0 = x_g = x$ and $y_0 = y_g = y$.

We use the inclusive fitness effects above to determine when selection favours the evolution of cooperation and cognition by looking at when $H > 0$ and $K > 0$, respectively. We are mainly interested in the cases where a population playing full cooperation and full cognition (i.e., $x^*, y^* = 1$) is stable. Hence, the conditions for the extreme point $x, y = 1$ to be convergence stable are if both $H(1,1) > 0$ and $K(1,1) > 0$. We later confirm these stability conditions, using individual-based simulations (Supplementary material; Figs. S6-8).

Our expression for Hamilton's selection gradient is in terms of 'whole-group' relatedness which includes relatedness to self. However, this measure of relatedness can also be expressed in terms of group size and 'others-only' relatedness R_o (Pepper, 2000), which measures the relatedness of a focal individual to a random member in the group (excluding the focal). By substituting $R = (1 + (n-1)R_o)/n$ into $H(x,y)$ and $K(x,y)$, we can determine when full cooperation and full cognition remain stable when varying both group size n

and R_o.

3 Results

By substituting the benefit and fitness functions into $H(x, y)$ and $K(x, y)$, we find that, assuming the population average fecundity is always positive, the selection gradients for cooperation $H(x, y)$ and cognition $K(x, y)$ are positive if

$$[1 + (n-1)R_o]\left(\frac{b}{n}(1-y) + y\frac{b+b_C}{n}\mathrm{e}^{\alpha y}\right) > c \tag{1a}$$

$$x[1 + (n-1)R_o]\left(\frac{b+b_C}{n}\mathrm{e}^{\alpha y}(1+\alpha y) - \frac{b}{n}\right) > d, \tag{1b}$$

respectively. Both selection gradients comprise the marginal returns from cooperation and cognition (last term on the left hand side in Ineq.(1)), weighted by relatedness of the individual to itself (i.e., 1) and to the other $n-1$ group members. Table 1 provides a summary of the conditions for full cooperation and full cognition ($x, y = 1$) to be favoured. Table S1 gives the same conditions in the absence of relatedness ($R_o = 0$).

In the Supplementary Information (§1), we consider a more explicit life cycle, where relatedness within patches emerges as a consequence of demographic processes, such as the probability of adult survival s and juvenile dispersal m Taylor (1992); Taylor and Irwin (2000). We find that the conditions for

cooperation and cognition to be favoured remain the same as in Ineq.(1) and Table (1), except that R_o is replaced by $\kappa = \{2(1-m)s\}/\{n[2-m(1-s)]+2(1-m)s\}$, where κ is the scaled relatedness coefficient which is demographically scaled so as to capture the effect of both increased genetic assortment and increased local competition between kin (Lehmann and Rousset, 2010). Our value of κ recovers that presented in Lehmann and Rousset (2010) for this life-cycle.

3.1 The evolution of cooperation

We first consider the evolution of cooperation, assuming no cognition in the population ($y \to 0$). In this case, cooperation will be favoured if, and only if $(n-1)R_o b/n > c - b/n$ (Fig.2a; Fig.S1). That is, if the benefits received from the proportion of relatives among the $n-1$ group members are greater than the net cost of contributing (i.e., cost of contribution minus own share). In other words, the indirect fitness benefits need to outweigh the direct fitness cost. This condition is Hamilton's Rule for the linear public goods game (Bijma and Aanen, 2010; Ohtsuki, 2014). Selection for cooperation is frequency-independent, and so the population will always converge towards full cooperation if Hamilton's Rule is satisfied. However, assuming $b/n < c$, cooperation can never be favoured in the absence of relatedness, in which case the population will converge towards full defection.

3.2 The evolution of cognition

If there is a sufficient level of cooperation in the population, then cognition can be favoured. Specifically, if $x^* \geq d/\{[1 + (n-1)R_o]b_C/n\}$ (Fig.2b). Here, both larger d and n increase the required amount of cooperation for cognition to be favoured, while larger b_C and R_o decrease it. Assuming that cooperation has fully invaded the population ($x = 1$), then cognition will invade if $[1 + (n-1)R_o b_C/n] > d$. The initial invasion of cognition does not depend synergy, because cognition can only provide synergistic returns if there is already some cognition in the population (i.e., $y > 0$). Consequently, if cognition only allows synergy, but no larger additive returns, then cognition can never increase from rarity (i.e., $K(x, 0) < 0$ if $d > 0$ and $b_C = 0$). However, this follows from our assumption of weak selection with continuous phenotypes, and would not necessarily occur with discrete phenotypes and strong selection.

3.3 The coevolution of cooperation and cognition

If cognition is favoured and increases in the population, it will in turn increase the selection pressure on cooperation, and vice versa. This is because, as the level of cognition (cooperation) increases in the population, the cognitively enhanced cooperators benefit increasingly from both their own contribution and that of their relatives. This can be seen from Ineq.(1), where H and K are increasing in y and x, respectively. The population will then converge towards full cooperation and full cognition (i.e., $x, y = 1$).

To summarise, cooperation can only evolve if there is sufficient relatedness R_o and benefits b (for a given c), whereas cognition can only evolve if there is sufficient cooperation (x^*), relatedness and benefits from cognition b_C (for a given d). In all cases, increasing group size n hinders the evolution of both traits. Therefore, it follows that the population can only end up in one of three different states: (i) full defection and no cognition, (ii) full cooperation and no cognition, and (iii) full cooperation and cognition.

The analytical conditions for full cooperation and cognition to be stable are summarised in Table 1. Full cooperation with full cognition is favoured by increasing the additional gains allowed by cognition (b_C), synergy (α) and relatedness (R_o) but disfavoured by increasing group size n and the cost of either cooperation c or cognition d (Fig.2a). In the absence of synergy, only the stability of cooperation, but not that of cognition, increases with increasing baseline benefit b. In contrast, with synergy only, increasing b increases the stability of both traits.

3.4 The transition from high to low relatedness

Assuming that a population has converged towards full cooperation and cognition, how stable would this population be in case of a subsequent decrease in relatedness? We can answer this by substituting $R_o = 0$ into the conditions shown in Table 1 which gives the results in Table S1.

The results depend upon the type of benefits provided by cognition. If cognition only allows for larger benefits from cooperation, then cooperation is

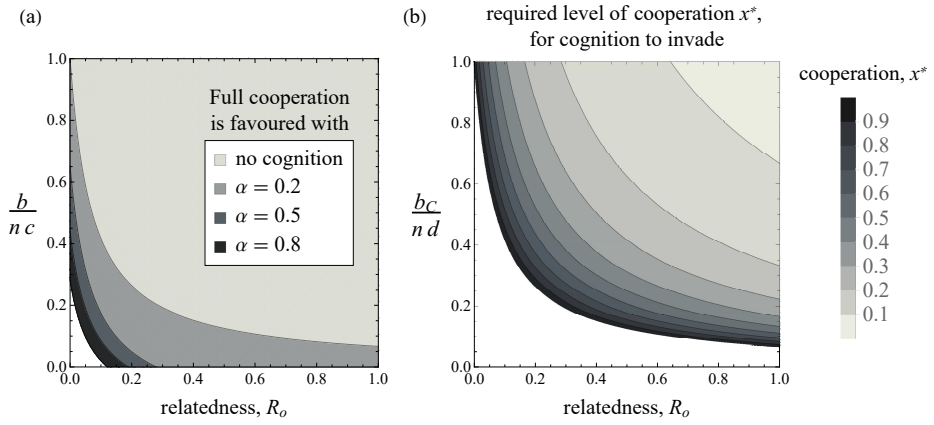

Figure 2: Invasion and stability conditions for cooperation and cognition. (a) Parameter space where full cooperation is stable (and can invade), with and without cognition, as a function of relatedness and the per capita benefit to cost ratio from cooperation (scaled by group size n). Increasing the synergy factor α increases the stability of full cooperation, even without relatedness ($R_o = 0$). (b) Level of cooperation required for cognition to invade, in function of relatedness and the per capita share to cost ratio from cognition (scaled by group size n). Parameters: in (a) $b_C = 2.5$; in both panels $n = 15$, $c = 0.1$, $d = 0.05$.

stable if the share from the public good $(b + b_C)/n$ exceeds the cost of contributing c, whereas cognition is stable if the share from the benefit of cognition (b_C/n) exceeds the cost of cognition d.

In contrast, if cognition only allows for synergy ($b_C = 0$ and $\alpha > 0$) then cooperation and cognition can be stable under less stringent conditions, and increasing synergy increases their stability. Finally, if cognition allows for both larger gains and synergy between cognitively enhanced individuals ($b_C > 0$ and $\alpha > 0$), then full cooperation and full cognition are stable over an even wider range of parameter space (Fig.2a; Fig.S2-5). As before, increasing patch size n disfavours cooperation and cognition, such that there is a threshold patch size above which they will not be stable (Fig.S2-5). In our more

explicit life-cycle, selection for cooperation and cognition decreases if the scaled relatedness coefficient κ decreases. This occurs if migration m and patch size n increase, or if survival s decreases (Supplementary information §1).

b_C	α	$H > 0$	$K > 0$
$= 0$	$= 0$	$\frac{b}{n}[1 + (n-1)R_o] > c$	$0 > d$
> 0	$= 0$	$\frac{b + b_C}{n}[1 + (n-1)R_o] > c$	$\frac{b_C}{n}[1 + (n-1)R_o] > d$
$= 0$	> 0	$\mathrm{e}^{\alpha}\frac{b}{n}[1 + (n-1)R_o] > c$	$[\mathrm{e}^{\alpha}(1+\alpha)\frac{b}{n} - \frac{b}{n}][1 + (n-1)R_o] > d$
> 0	> 0	$\mathrm{e}^{\alpha}\frac{b + b_C}{n}[1 + (n-1)R_o] > c$	$[\mathrm{e}^{\alpha}(1+\alpha)\frac{b + b_C}{n} - \frac{b}{n}][1 + (n-1)R_o] > d$

Table 1: Conditions for full cooperation and cognition to be favoured (i..e, $H(1,1) > 0$ and $K(1,1) > 0$) for different scenarios. Parameters: R_o, relatedness; c, cooperation cost; d, cognition cost; b, cooperation benefit; b_C, cognition benefit; α, synergy factor; n, group size.

3.5 Computer simulations

We use individual-based simulations to confirm whether a population at the equilibrium can remain stable even as relatedness decreases. We give a detailed description of the simulation model in the supplementary material (§2). We start our simulations with different levels of relatedness, and halfway through, we reduce relatedness to a value of $R_o = 1/100$. We then check the long-term average of both cooperation and cognition in the population. All simulations were run for 10^6 generations.

As predicted by our analytical model, full cooperation and full cognition remained locked even after a drop in relatedness, provided sufficient synergy and additive gains from cognition (Fig.3, Fig.S7-8). In the absence of cognition and whenever the direct benefits from cooperation and cognition were too low, populations which initially evolved cooperation were invaded by defectors as soon as relatedness decreased to negligible values. We confirmed these results for a large range of parameter values (Fig.S6-8).

3.6 Machiavellian cognition

So far, we have assumed that cognition was beneficial for all group members, because the presence of cognitively enhanced individuals increased the share received by each of their patch members. However, it has been argued that cognition could be used for selfish profits through manipulation and deception of social partners (the 'Machiavellian Intelligence hypothesis'; Byrne and Whiten (1989); Whiten and Byrne (1997)). In this section, we explore how

this negative aspect of cognition affects its coevolution with cooperation.

We expand our baseline model by assuming that relatively more intelligent individuals can exploit their partners by taking more than their fair share of the public good, e.g. through manipulation. Specifically, the benefit received by a focal individual with cognition y_0 is now $B(x_g, y_g)(1 + gy_0)/(1 + gy_g)$. The parameter g controls the magnitude of exploitation in the following way: if $g \to 0$ or if $y_0 = y_g$, a focal individual will receive the same public good share as everyone else. If $y_0 \neq y_g$ and $g \to \infty$, the focal individual will receive a proportion y_0/y_g of the per capita benefit from the public good. Here, the cognition trait is analogous to the exploitation trait investigated in Brown and Taylor (2010), except that cognition can increase and, at the same time, decrease the public good benefit to others if $\alpha, b_C > 0$ and $g > 0$, respectively.

The selection pressure on cooperation does not change in this scenario (supplementary information §3). This comes from the fact that cognition still has the same effect on the public good created as in the baseline model. However, if there is cooperation in the population, the selection pressure for cognition is considerably larger than in our baseline model (Table S2, Fig. S9). This is because larger b and g also increase the marginal benefit from cognition. So, cognition can invade from rarity even in the absence of relatedness, if there is a sufficient level of cooperation in the population (i.e., $x^* \geq dn/[\, b_C + bg(n-1)]\,$). As a consequence, full cognition is also stable under a larger area of parameter space (Table S2). Full cognition can even be stable even in the absence of both larger benefits and synergism (b_C, $a = 0$).

This is not surprising, because in a population with full cooperation and full cognition, a focal mutant with a lower level of cognition will be exploited by its partners.

We also find that increasing relatedness can sometimes have a negative effect on the evolution of Machiavellian cognition (supplementary information §3). The reason is that exploiting partners leads to indirect fitness costs if too many partners are relatives. In the extreme case where cognitively enhanced individuals produce neither larger, nor synergistic benefits ($b_C, \alpha = 0$), higher relatedness always hinders the evolution of cognition (supplementary information §3).

4 Discussion

We found that coevolution between cooperation and cognition can lead to a transition in the nature of the cooperative dilemma—from a state where cooperation can only evolve and be maintained with sufficient relatedness, to a state where it is stable even without relatedness. In small groups of related individuals, cooperation can evolve (Fig.2b). If cognition enables individuals to either generate larger gains from cooperation and/or exploit less intelligent partners, then cognition can invade into cooperative societies (Fig.2b). When cooperation and cognition are both favoured, an increase in either trait will increase selection pressure for the other trait, leading to an evolutionary feedback loop until the population has converged toward full cooperation and cognition. At this point, if relatedness drops via an increase

in either group size or dispersal, such that cooperation on its own would not have been favoured, the combination of cooperation and cognition can still be stable. Cognition can therefore lead to a population being 'locked' in a stable cooperative state (Fig.2-3).

4.1 Cooperation and cognition

Cognition can only evolve in our model if there is a sufficient level of cooperation in the population (Fig.2b). We focused on relatedness as the underlying factor promoting the emergence of cooperation. That relatedness can often play a key role in the evolution of cooperation, in scenarios where it is then less important for its maintenance, as has also been argued with reciprocity, punishment, and group augmentation (Axelrod and Hamilton, 1981; Kokko et al., 2001; Gardner and West, 2004; West et al., 2007; André, 2015). However, we believe that our conclusion on how cognition could have transformed the type of dilemmas faced by early humans does not rely on relatedness. In fact, processes where an additional trait can transform the payoff matrix into a game where cooperation maximises selfish profits have been argued to be common in nature (Bshary et al., 2016). Cooperation could have been promoted if individuals within groups depended on each other's cooperation to survive against predators, environmental catastrophes, or other groups (the 'interdependence hypothesis'; Roberts (2005); Tomasello et al. (2012); Reader et al. (2011); Kokko et al. (2001); Keeley (1997); Gavrilets (2015)). In turn, cooperation is likely to have introduced selection pressures on individuals to benefit from cooperation even more, or at a reduced cost (Bshary

et al., 2016). For example, Gavrilets (2015) showed that cooperation and cognition could evolve without relatedness when groups compete with each other. Gavrilets investigated a different productivity function for the baseline public good, and cognition did not have any synergistic effects on cooperation. Therefore, cognition in his model did not change the nature of the cooperative dilemma in the same way as in our model, and so cooperation might not remain stable if inter-group conflict becomes less frequent.

We also found that, if cognition does not allow for exploitation of others, it could evolve only if cognitively enhanced individuals are able to generate larger amounts of public goods (Fig.2b). The kin benefits from cognition need to outweigh the cost of developing cognition, which implies sufficiently high relatedness between partners and gains from cognition. However, our assumption that cognition provides benefits only through group cooperation is restrictive. In nature, cognition might have provided benefits in various additional contexts, such as tool making or foraging, thereby favouring its evolution (Rosati, 2017).

Furthermore, we found that synergy is not important for the invasion of cognition. This is because cognition cannot provide significant synergistic returns to rare mutants, unless there is already some level of cognition in the population. On the other hand, synergy is crucial for the stability of both traits in the absence of relatedness (Fig.2a; Fig.3). Biologically, synergy could occur if cognition allows high coordination between hunters, via the use of planning and sign language, or acting in function of the other hunters' movement (Boesch, 2002). The stabilising effect of synergism on cooperation

is well-known (Motro, 1991; Hauert et al., 2006; Archetti and Scheuring, 2011, 2012; Ohtsuki, 2012; Peña et al., 2015). However, most previous studies investigating non-linear production functions focused on a single cooperative trait (although see Brown and Taylor (2010)). In contrast, we separated synergistic cooperation into two different traits. This allowed us to determine how both cooperation and synergistic cognition alter the selection pressure on each other.

4.2 Machiavellian cognition

We have also shown that cognition initially evolves and remains stable more readily if it allows the exploitation of less intelligent partners (Machiavellian intelligence hypothesis; Byrne and Whiten (1989); Whiten and Byrne (1997)). This is in line with previous studies which found that cooperation creates selection pressures for higher cognitive abilities leading to the deception and manipulation of others (McNally et al., 2012; McNally and Jackson, 2013). Our scenario is also very similar to that in Brown and Taylor (2010), where cooperation coevolves with an exploitative trait that reduces the amount of public good for personal profit. Our model complements these studies as we have incorporated both the beneficial and harmful consequences of cognition. This revealed that both Machiavellian and beneficial cognition can evolve together, provided the beneficial effect of cognition on the public good is sufficiently large. This suggests that, in accordance with the 'social intelligence hypothesis', cognition could have evolved due to its various effects on social interactions (Reader et al., 2011; Dunbar, 1998; Byrne and

Whiten, 1989; Whiten and Byrne, 1997; Tomasello et al., 2005). On the other hand, we showed that in the absence of greater benefits from cognition ($b_C, \alpha = 0$), Machiavellian cognition evolves more easily in groups with low, rather than high relatedness, as was previously suggested (McNally et al., 2012).

4.3 Key predictions

Our model can be validated by either testing our assumptions or predictions. First, one of our assumptions was that cognition rapidly increases the benefits (or decreases the costs) of cooperation, i.e., there is synergy between cognitively enhanced individuals. Because it is difficult to determine cognition in real world collective actions, a starting point would be to estimate it indirectly. One proxy for cognition would be the level of specialisation (skills acquired through learning) required for hunting parties in modern hunter-gathers. Hence, demonstrating synergy requires showing that, for example, the success rate of groups with ten specialised hunters is more than twice that of groups with five specialised hunters. An important point here is to control for observability, because hunters in smaller groups might be more isolated from each other, thereby providing more opportunities for cheating. Specialised hunters might also simply be better at detecting cheats. Second, our prediction that cognition allows for the maintenance of cooperation in the absence of other promoting factors can be tested in laboratory experiments, by artificially manipulating cognition. A possible experiment would be to recreate a situation similar to that in our Fig.3. For example, in a

cooperative task where coordination (or learning) provides larger benefits, cooperation could be initially promoted (e.g. through global competition West et al. (2006)). Then, the cooperation-promoting mechanism could be removed halfway through. We would then expect cooperation to remain stable in a treatment where coordination/learning is allowed, compared to when it is not allowed.

Third, another prediction from our model is that cognition allows stable cooperation levels without relatedness or any enforcement mechanism, such as reciprocity, partner choice or punishment (West et al., 2007). Indeed, enforcement mechanisms usually require cognition and can also maintain cooperation without relatedness (West et al., 2007; Bshary et al., 2016). Hence, one way to validate our prediction would be to compare relatedness between social partners across different cooperative tasks within primate species. We would then expect lower relatedness between social partners and the absence of partner control in those cooperative tasks that are more cognitively demanding. Finally, a more general prediction is that we expect more cooperation in more intelligent species. This could be tested with comparative analyses on different primate species, by looking at the correlation between between cooperation and relatedness, and including cognition as a covariate. This is already partially supported by the positive correlation between cooperation and deception observed in primates (McNally and Jackson, 2013). As before, an important point here will be to distinguish between cases of cooperation with and without partner control mechanisms.

Ethics

No organisms/human participants were used in this study.

Data accessibility

The simulation code and data are available from the OSF data repository (osf.io/7p6us).

Authors Contributions

M.D.S and S.A.W. conceived the study and contributed to modelling. M.D.S. performed the simulations. M.D.S. and S.A.W. wrote the paper. All authors gave their final approval for publication.

Competing interests

We have no competing interests.

Funding

M.D.S. is funded by the Swiss National Science Foundation (grant P2LAP3-158669).

Acknowledgments

We thank Guy Cooper, Asher Leeks, Sam Levin, Matishalin Patel, Tom Scott and Peter Taylor for comments and/or discussion. We are grateful to Nichola Raihani for comments and for inviting us to contribute to this issue. We would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work. http://dx.doi.org/10.5281/zenodo.22558.

References

Aiello, L. C. and R. I. Dunbar (1993): "Neocortex size, group size, and the evolution of language," *Current anthropology*, 34, 184–193.

Alexander, R. D. (1990): "How did humans evolve? Reflections on the uniquely unique species." .

André, J.-B. (2015): "Contingency in the Evolutionary Emergence of Reciprocal Cooperation." *The American Naturalist*, 185, 303–316.

Andres, A. M., M. Soldevila, A. Navarro, K. K. Kidd, B. Oliva, and J. Bertranpetit (2004): "Positive selection in MAOA gene is human exclusive: determination of the putative amino acid change selected in the human lineage," *Human genetics*, 115, 377–386.

Archetti, M. and I. Scheuring (2011): "Coexistence of cooperation and defection in public goods games," *Evolution*, 65, 1140–1148.

——— (2012): "Review: Game theory of public goods in one-shot social

dilemmas without assortment," *Journal of Theoretical Biology*, 299, 9 – 20, evolution of Cooperation.

Axelrod, R. and W. Hamilton (1981): "The evolution of cooperation," *Science*, 211, 1390–1396.

Barker, G. (2009): *The agricultural revolution in prehistory: why did foragers become farmers?*, Oxford University Press on Demand.

Bijma, P. and D. K. Aanen (2010): "Assortment, Hamilton's rule and multilevel selection," *Proceedings of the Royal Society of London B: Biological Sciences*, 277, 673–675.

Bittles, A. H. and M. Black (2010): "Consanguineous marriage and human evolution," *Annual Review of Anthropology*, 39, 193–207.

Bocquet-Appel, J.-P. (2011): "When the world's population took off: the springboard of the Neolithic demographic transition," *Science*, 333, 560–561.

Boesch, C. (2002): "Cooperative hunting roles among taï chimpanzees," *Human Nature*, 13, 27–46.

Brosnan, S. F., L. Salwiczek, and R. Bshary (2010): "The interplay of cognition and cooperation," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 2699–2710.

Brown, S. P. and P. D. Taylor (2010): "Joint evolution of multiple social traits: a kin selection analysis," *Proceedings of the Royal Society of London B: Biological Sciences*, 277, 415–422.

493 Bshary, R., K. Zuberbühler, and C. P. van Schaik (2016): "Why
494 mutual helping in most natural systems is neither conflict-free nor based
495 on maximal conflict," *Phil. Trans. R. Soc. B*, 371, 20150091.

496 Buss, D. (2015): *Evolutionary psychology: The new science of the mind*,
497 Psychology Press.

498 Byrne, R. and A. Whiten (1989): "Machiavellian intelligence: social
499 expertise and the evolution of intellect in monkeys, apes, and humans
500 (oxford science publications)," .

501 Call, J. (2009): "Contrasting the Social Cognition of Humans and Non-
502 human Apes: The Shared Intentionality Hypothesis," *Topics in Cognitive*
503 *Science*, 1, 368–379.

504 Dunbar, R. I. (2009): "The social brain hypothesis and its implications for
505 social evolution," *Annals of human biology*, 36, 562–572.

506 Dunbar, R. I. M. (1998): "The social brain hypothesis," *Evolutionary*
507 *Anthropology*, 6, 178–190.

508 Ferland, R. J., W. Eyaid, R. V. Collura, L. D. Tully, R. S.
509 Hill, D. Al-Nouri, A. Al-Rumayyan, M. Topcu, G. Gascon,
510 A. Bodell, et al. (2004): "Abnormal cerebellar development and ax-
511 onal decussation due to mutations in AHI1 in Joubert syndrome," *Nature*
512 *genetics*, 36, 1008–1013.

513 Frank, S. A. (1998): *Foundations of social evolution*, Princeton University
514 Press.

Gardner, A. and S. A. West (2004): "Cooperation and punishment, especially in humans," *The American Naturalist*, 164, 753–764.

Gavrilets, S. (2015): "Collective action and the collaborative brain," *Journal of The Royal Society Interface*, 12, 20141067.

Gazda, S. K., R. C. Connor, R. K. Edgar, and F. Cox (2005): "A division of labour with role specialization in group–hunting bottlenose dolphins (Tursiops truncatus) off Cedar Key, Florida," *Proceedings of the Royal Society of London B: Biological Sciences*, 272, 135–140.

Hamilton, W. D. (1964): "The genetical evolution of social behaviour. II," *Theoretical Population Biology*, 7, 17 – 52.

Hatala, K. G., N. T. Roach, K. R. Ostrofsky, R. E. Wunderlich, H. L. Dingwall, B. A. Villmoare, D. J. Green, J. W. Harris, D. R. Braun, and B. G. Richmond (2016): "Footprints reveal direct evidence of group behavior and locomotion in Homo erectus," *Scientific reports*, 6, 28766.

Hauert, C., F. Michor, M. A. Nowak, and M. Doebeli (2006): "Synergy and discounting of cooperation in social dilemmas," *Journal of Theoretical Biology*, 239, 195 – 202.

Hill, K. R., R. S. Walker, M. Božičević, J. Eder, T. Headland, B. Hewlett, A. M. Hurtado, F. Marlowe, P. Wiessner, and B. Wood (2011): "Co-Residence Patterns in Hunter-Gatherer Societies Show Unique Human Social Structure," *Science*, 331, 1286–1289.

Keeley, L. H. (1997): *War before civilization*, OUP USA.

Kokko, H., R. A. Johnstone, and C.-B. T. H. (2001): "The evolution of cooperative breeding through group augmentation," *Proceedings of the Royal Society of London B: Biological Sciences*, 268, 187–196.

Krause, J., C. Lalueza-Fox, L. Orlando, W. Enard, R. E. Green, H. A. Burbano, J.-J. Hublin, C. Hänni, J. Fortea, M. De La Rasilla, et al. (2007): "The derived FOXP2 variant of modern humans was shared with Neandertals," *Current biology*, 17, 1908–1912.

Lalueza-Fox, C., A. Rosas, A. Estalrrich, E. Gigli, P. F. Campos, A. Garca-Tabernero, S. Garca-Vargas, F. Snchez-Quinto, O. Ramrez, S. Civit, M. Bastir, R. Huguet, D. Santamara, M. T. P. Gilbert, E. Willerslev, and M. de la Rasilla (2011): "Genetic evidence for patrilocal mating behavior among Neandertal groups," *Proceedings of the National Academy of Sciences*, 108, 250–253.

Lehmann, L. and F. Rousset (2010): "How life history and demography promote or inhibit the evolution of helping behaviours," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365, 2599–2617.

Marlowe, F. W. (2005): "Hunter-gatherers and human evolution," *Evolutionary Anthropology: Issues, News, and Reviews*, 14, 54–67.

McNally, L., S. P. Brown, and A. L. Jackson (2012): "Cooperation and the evolution of intelligence," *Proceedings of the Royal Society of*

London B: Biological Sciences.

McNally, L. and A. L. Jackson (2013): "Cooperation creates selection for tactical deception," in *Proc. R. Soc. B*, The Royal Society, vol. 280, 20130699.

Motro, U. (1991): "Co-operation and defection: Playing the field and the ESS," *Journal of Theoretical Biology*, 151, 145 – 154.

Ohtsuki, H. (2012): "Does synergy rescue the evolution of cooperation? An analysis for homogeneous populations with non-overlapping generations," *Journal of theoretical biology*, 307, 20–28.

——— (2014): "Evolutionary dynamics of n-player games played by relatives," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369, 20130359.

Peña, J., G. Nöldeke, and L. Lehmann (2015): "Evolutionary dynamics of collective action in spatially structured populations," *Journal of theoretical biology*, 382, 122–136.

Pepper, J. W. (2000): "Relatedness in trait group models of social evolution," *Journal of Theoretical Biology*, 206, 355–368.

Pruetz, J. D. and P. Bertolani (2007): "Savanna Chimpanzees, Pan troglodytes verus, Hunt with Tools," *Current Biology*, 17, 412 – 417.

Reader, S. M., Y. Hager, and K. N. Laland (2011): "The evolution of primate general and cultural intelligence," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 366, 1017–1027.

583 Rightmire, G. P. (2004): "Brain size and encephalization in early to Mid-
584 Pleistocene Homo," *American Journal of Physical Anthropology*, 124, 109–
585 123.

586 Roberts, G. (2005): "Cooperation through interdependence," *Animal Be-*
587 *haviour*, 70, 901–908.

588 Rosati, A. G. (2017): "Foraging cognition: reviving the ecological intelli-
589 gence hypothesis," *Trends in cognitive sciences*, 21, 691–702.

590 Schoenemann, P. T. (2006): "Evolution of the size and functional areas
591 of the human brain," *Annu. Rev. Anthropol.*, 35, 379–406.

592 Smith, T. G., D. B. Siniff, R. Reichle, and S. Stone (1981): "Coor-
593 dinated behavior of killer whales, Orcinus orca, hunting a crabeater seal,
594 Lobodon carcinophagus," *Canadian Journal of Zoology*, 59, 1185–1189.

595 Somel, M., X. Liu, and P. Khaitovich (2013): "Human brain evolu-
596 tion: transcripts, metabolites and their regulators," *Nature Reviews Neu-*
597 *roscience*, 14, 112–127.

598 Stout, D., E. Hecht, N. Khreisheh, B. Bradley, and T. Chami-
599 nade (2015): "Cognitive demands of Lower Paleolithic toolmaking," *PLoS*
600 *One*, 10, e0121804.

601 Stout, D., N. Toth, K. Schick, and T. Chaminade (2008): "Neural
602 correlates of Early Stone Age toolmaking: technology, language and cogni-
603 tion in human evolution," *Philosophical Transactions of the Royal Society*
604 *of London B: Biological Sciences*, 363, 1939–1949.

Taylor, P. D. (1992): "Altruism in viscous populationsan inclusive fitness model," *Evolutionary ecology*, 6, 352–356.

Taylor, P. D. and S. A. Frank (1996): "How to Make a Kin Selection Model," *Journal of Theoretical Biology*, 180, 27 – 37.

Taylor, P. D. and A. J. Irwin (2000): "Overlapping generations can promote altruistic behavior," *Evolution*, 54, 1135–1141.

Tomasello, M., M. Carpenter, J. Call, T. Behne, and H. Moll (2005): "Understanding and sharing intentions: The origins of cultural cognition," *Behavioral and brain sciences*, 28, 721–727.

Tomasello, M., A. P. Melis, C. Tennie, E. Wyman, and E. Herrmann (2012): "Two Key Steps in the Evolution of Human Cooperation: The Interdependence Hypothesis," *Current Anthropology*, 53, 673–692.

Vallender, E. J., N. Mekel-Bobrov, and B. T. Lahn (2008): "Genetic basis of human brain evolution," *Trends in Neurosciences*, 31, 637 – 644.

West, S. A., A. Gardner, D. M. Shuker, T. Reynolds, M. Burton-Chellow, E. M. Sykes, M. A. Guinnee, and A. S. Griffin (2006): "Cooperation and the scale of competition in humans," *Current Biology*, 16, 1103–1106.

West, S. A., A. S. Griffin, and A. Gardner (2007): "Evolutionary explanations for cooperation," *Current Biology*, 17, R661–R672.

626 Whiten, A. and R. Byrne (1997): "Machiavellian intelligence 2: evalua-
627 tions and extensions," .

628 Whiten, A. and D. Erdal (2012): "The human socio-cognitive niche and
629 its evolutionary origins," *Philosophical Transactions of the Royal Society*
630 *of London B: Biological Sciences*, 367, 2119–2129.

631 Wright, S. (1931): "Evolution in mendelian populations," *Genetics*, 16,
632 97–159.

633 Yeshurun, R., G. Bar-Oz, and M. Weinstein-Evron (2007): "Mod-
634 ern hunting behavior in the early Middle Paleolithic: Faunal remains from
635 Misliya Cave, Mount Carmel, Israel," *Journal of Human Evolution*, 53,
636 656 – 677.

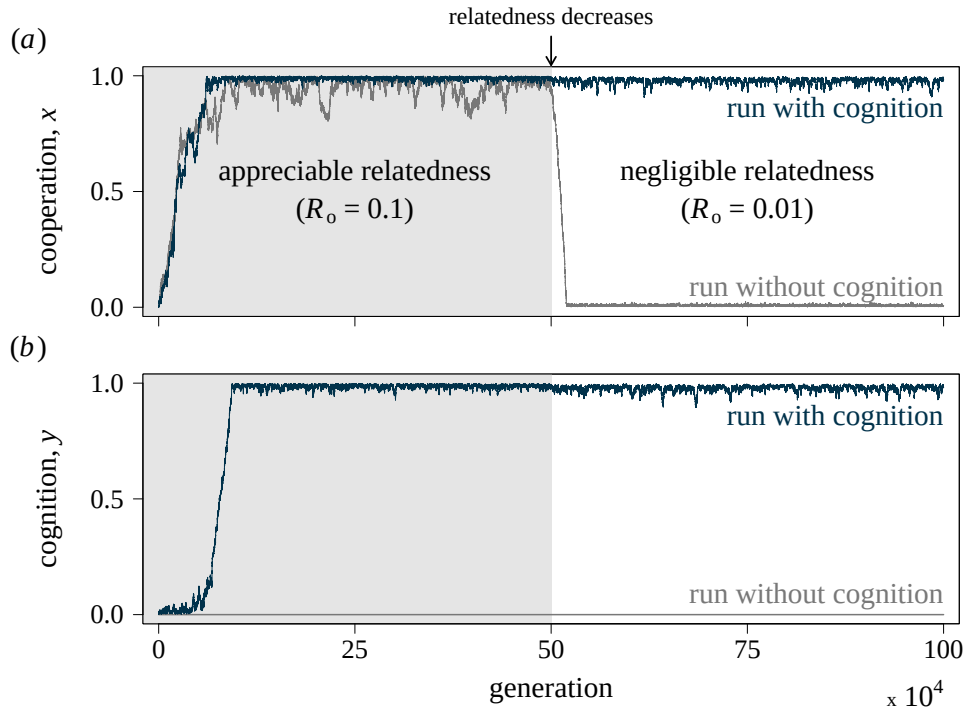

Figure 3: Joint evolution of cooperation and cognition in individual-based simulations. We compare the stability of cooperation (a) and cognition (b), with and without cognition. In both panels, the blue (grey) line represents the same simulation run with (without) cognition. In the run with cognition, the synergy factor $\alpha = 0.4$. Relatedness was decreased half-way through each run (i.e., $k = 10$ and $k = 100$ in the first and second half, respectively, with $R_o = 1/k$). Parameters: $n = 15$, $c = 1$, $b = 7$, $d = 0.5$, $b_C = 3.5$, $\mu_h = \mu_c = 0.01$, $\sigma = 0.01$.